Work extremum principle: Structure and function of quantum heat engines

Armen E. Allahverdyan

Yerevan Physics Institute, Alikhanian Brothers Street 2, Yerevan 375036, Armenia

Ramandeep S. Johal

Post-Graduate Department of Physics, Lyallpur Khalsa College, Jalandhar 144001, Punjab, India

Guenter Mahler

Institute of Theoretical Physics I, University of Stuttgart, Pfaffenwaldring 57, 70550 Stuttgart, Germany
(Dated: February 2, 2008)

We consider a class of quantum heat engines consisting of two subsystems interacting via a unitary transformation and coupled to two separate baths at different temperatures $T_h > T_c$. The purpose of the engine is to extract work due to the temperature difference. Its dynamics is not restricted to the near equilibrium regime. The engine structure is determined by maximizing the extracted work under various constraints. When this maximization is carried out at finite power, the engine dynamics is described by well-defined temperatures and satisfies the local version of the second law. In addition, its efficiency is bounded from below by the Curzon-Ahlborn value $1 - \sqrt{T_c/T_h}$ and from above by the Carnot value $1 - (T_c/T_h)$. The latter is reached—at finite power—for a macroscopic engine, while the former is achieved in the equilibrium limit $T_h \to T_c$. When the work is maximized at a zero power, even a small (few-level) engine extracts work right at the Carnot efficiency.

PACS numbers: PACS: 05.30.-d, 05.70.Ln

I. INTRODUCTION.

Heat engines are natural or artificial devices, the goal of which is to extract work (high-graded energy) from non-equilibrium sources of heat (low-graded energy) [1, 2]. There are *three* basic characteristics of the engine operation: *i)* the work extracted per cycle; *ii)* the efficiency with which the input heat is converted into work; *iii)* the power of work-extraction, i.e. the work-extracted per cycle divided over the cycle duration.

Heat engines may evolve in time toward states that provide (constrained) optimization of their functional characteristics. Here are three examples from largely different fields:

1. The modern artificial engines are much more powerful and efficient than those which started the Industrial Revolution. The cause of this improvement are engineering efforts driven by our desire to get more high-graded energy at a lower cost.

2. A similar trend is seen in biology and ecology. Driven by evolution, higher organisms and more developed ecosystems have more refined and more optimal means of extracting energy from their respective environment [3]. This observation led to several quantitative formulations, which were verified experimentally [3].

3. The earth atmosphere can approximately be regarded as a huge heat-engine operating between two thermal baths (cold bath of the earth surface and eventually the hot bath of the sun) and producing as output large-scale turbulent motion of air and vapor [4]. As verified by observation [4], this engine is also tuned to extract the maximal work [4], although the precise mechanism of this tuning is unclear yet.

Needless to stress that in all above situation the op-timization of heat engines proceeds in the presence of constraints that determine the very path of the engine evolution. This fact is especially important in the above bio-ecological perspective.

The purpose of this paper is to understand the structure of quantum heat engines emerging from the maximization of work (produced per cycle) under specific constraints. The study of quantum engines started in 60's [5], when it was realized that many models of lasers and masers are in fact quantum heat engines. A good review of this early activity is given in [6]. Nowadays the physics of quantum heat engines is a rich field[7–9, 11–21] related to other branches of modern quantum theory; see [22] for a recent review of engines in the context of quantum information theory. We see two basic reasons for studying quantum heat engines: *i)* understanding of how thermodynamics emerges from the quantum mechanics; *ii)* clarifying principal possibilities of nano-scale devices.

Our analysis starts from quantum mechanics and does not rely on the validity of thermodynamical concepts normally invoked in studying heat engines [1, 2]. (Relying on quantum mechanics does not preclude us from considering macroscopic systems.) We allow the intermediate states of the engine to be arbitrary far from equilibrium. We shall however see that local thermodynamical concepts—such as the existence of local temperatures in the intermediate stages of the engine functioning, or the validity of local formulations of the second law—emerge as a result of maximizing the produced work.

Our model will consist of two quantum systems R and S interacting with thermal baths at temperatures T_h and T_c, respectively $(T_h > T_c)$; see Fig. 1. The number of energy levels for R and S is finite, but it can be made very large going to the macroscopic limit. R and S interact

with an external work-source producing work and then relax back to equilibrium under influence of the baths. For given temperatures (and given number of subsystem energy levels) work maximization can be introduced on three different levels: (1.) One optimizes the extracted work over the interaction of R and S with the external source of work (2.) One maximizes the work, in addition, over the spectral structure of R and S. (3.) Finally, one optimizes the extracted work also over the interaction of R and S with their respective thermal baths. In the following we will investigate these strategies in full detail.

Our main results are that on the second level of optimization the efficiency at the maximal work is always bounded from below by the Curzon-Ahlborn value $1 - \sqrt{T_c/T_h}$. This is in addition to the upper bound given by the Carnot value $1 - (T_c/T_h)$, which for the present setup holds generally, i.e., without any optimization. For any number of energy levels of R and S the Curzon-Ahlborn value is reached in the equilibrium limit $T_h \to T_c$. When the number of levels for R and S become very large (macroscopic limit), the efficiency reaches the Carnot value $1 - (T_c/T_h)$. In contrast to the standard thermodynamic realization of the Carnot efficiency, in our situation this efficiency is reached at a finite power. The third level of optimization is related to making the power very small. Now even a finite (two)-level systems R and S may be employed for extraction of a finite [per cycle] amount of work right at the Carnot efficiency.

This paper is organized as follows. Section II describes the studied model for a quantum heat engine. Here we also recall the derivation of the Carnot bound for the efficiency and address the problem of the power of work. In section III we discuss the optimization of work over the interaction with the sources. In section IV we report the results emerged from maximizing the work over the spectral structure of the involved quantum systems. In particular, we show that the Curzon-Ahlborn value is a lower bound for the efficiency, and we describe how one can employ a macroscopic engine for work-extraction at a finite power and with the Carnot efficiency. In section V we sacrify the finite-power condition, and show that a finite-level quantum engine can extract work right at the Carnot efficiency. The last section presents our conclusions, discusses some open questions and compares our findings with the results obtained previously. Several technical issues are discussed in appendices.

II. THE MODEL

Our engine consists of two quantum systems R and S which individually and in parallel couple to two different thermal baths and interact with a source of work; see Fig. 1. In the first step R and S undergo a work extraction process driven by the source. In the second step R and S relax back to their initial states under influence of the baths. Thus thermodynamic machine cycle is performed.
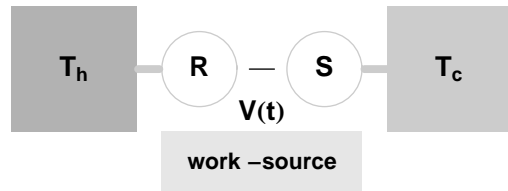

FIG. 1: A schematic representation of the considered quantum engine. Quantum systems R and S interact with thermal baths at temperatures T_h and T_c, respectively. The mutual interaction between them is governed by the potential $V(t)$ that also serves to deliver work to an external source. The mutual interaction is switched on for a short time only. Once it is switched off, the systems S and R do relax to equilibrium under influence of the respective thermal baths. Similar constructions of quantum heat engines appeared in [7, 10, 11].

A. Initialization

Consider the two finite quantum systems R and S with Hamiltonians H_R and H_S, respectively. R (S) has n (m) energy levels. In the scenarios studied here almost all operators involved will be diagonal in the energy representation. Thus we write

$$H_R = \text{diag}[\varepsilon_n, ..., \varepsilon_1], \quad \varepsilon_n \geq ... \geq \varepsilon_1, \qquad (1)$$
$$H_S = \text{diag}[\mu_m, ..., \mu_1], \quad \mu_m \geq ... \geq \mu_1, \qquad (2)$$

where $\text{diag}[a, .., b]$ is a diagonal matrix with entries $(a, ..., b)$. We will take

$$\varepsilon_1 = \mu_1 = 0, \qquad (3)$$

that is the energies of both systems are measured from zero. It will be seen below that this choice does not imply any loss of generality. Thus, ε_k and μ_l are the energy level spacings. The total Hamiltonian is

$$H = H_R \otimes 1 + 1 \otimes H_S = \sum_{\alpha=1}^{nm} E_\alpha |E_\alpha\rangle\langle E_\alpha|, \qquad (4)$$

$$= \text{diag}[\varepsilon_n + \mu_n, \varepsilon_n + \mu_{n-1}, ..., \varepsilon_1 + \mu_2, \varepsilon_1 + \mu_1], (5)$$

where $\{E_\alpha\}_{\alpha=1}^{nm}$ and $\{|E_\alpha\rangle\}_{\alpha=1}^{nm}$ are, respectively, eigenvalues and eigenvectors of $H_R \otimes 1 + 1 \otimes H_S$.

Subsystem R (S) is now assumed to couple to a thermal bath at temperature T_h (T_c); see Fig. 1. We shall assume:

$$T_h > T_c. \qquad (6)$$

At this stage there is no mutual coupling between R and S, and each bath drives its corresponding subsystem to an equilibrium state specified by the Gibbss density matrices

$$\rho = \frac{e^{-\beta_h H_R}}{\text{tr}\, e^{-\beta_h H_R}}, \quad \sigma = \frac{e^{-\beta_c H_S}}{\text{tr}\, e^{-\beta_c H_S}}, \qquad (7)$$

where $T_h = 1/\beta_h$ and $T_c = 1/\beta_c$. Alternatively,

$$\rho = \text{diag}[r_n, ..., r_1], \quad r_n \leq ... \leq r_1, \tag{8}$$

$$\sigma = \text{diag}[s_m, ..., s_1], \quad s_m \leq ... \leq s_1, \tag{9}$$

$$\varepsilon_k = T_h \ln \frac{r_1}{r_k}, \quad k = 1, ..., n, \tag{10}$$

$$\mu_l = T_c \ln \frac{s_1}{s_l}, \quad l = 1, ..., m. \tag{11}$$

The overall initial density matrix is

$$\Omega_{\text{in}} = \rho \otimes \sigma = \sum_{\alpha=1}^{nm} P_\alpha |P_\alpha\rangle\langle P_\alpha| \tag{12}$$

$$= \text{diag}[r_n s_n, r_n s_{n-1}, r_n s_{n-2}, ..., r_1 s_1], \tag{13}$$

and the average initial energy

$$\text{tr}[H \Omega_{\text{in}}] = \sum_{\alpha=1}^{nm} P_\alpha E_\alpha. \tag{14}$$

B. Two step process

1. First step: Unitary transformation

Now R + S are taken to interact via a time-dependent potential $V(t)$ so that their Hamiltonian reads

$$\mathcal{H}(t) = H_R \otimes 1 + 1 \otimes H_S + V(t), \tag{15}$$

$$V(t) = 0 \quad \text{for} \quad t < 0 \quad \text{and} \quad \text{for} \quad t > \tau. \tag{16}$$

This means that the initial and final Hamiltonian (i.e. before and after this first step) are both given by (4). The potential $V(t)$ in (15) is assumed to be sufficiently strong [and to act in a sufficiently short time] so that the influence of the thermal baths between the times 0 and τ can be neglected. We require this step to be thermally isolated, since, otherwise, for the couplings to the baths being on, one would get a direct heat exchange between the baths, a factor that should decrease the overall efficiency of the work-extraction. Thus the dynamics of R + S is unitary for $0 \leq t \leq \tau$:

$$\Omega_{\text{fin}} = \mathcal{U} \Omega_{\text{in}} \mathcal{U}^\dagger, \quad \mathcal{U} = \mathcal{T} \exp\left[-i \int_0^\tau d\bar{t}\, \mathcal{H}(\bar{t})\right], \tag{17}$$

where Ω_{in} is the initial density matrix defined in (12), Ω_{fin} is the final density matrix, $\mathcal{U}$ is the unitary evolution operator, and $\mathcal{T} \exp$ is the chronologically ordered exponent. The work extracted in this step reads

$$W_{nm}(T_h, T_c, \varepsilon, \mu, V) = \text{tr}[H(\Omega_{\text{in}} - \Omega_{\text{fin}})], \tag{18}$$

This work W depends on a set of parameters: The number of energy levels n and m for R and S, respectively, the spectra ε and μ according to (2), and the potential $V(t)$. In the following we do not make explicit the dependence of W on the temperatures T_h and T_c, which are taken to be fixed conditions in all cases.

2. Second step: Relaxation

After $V(t)$ has been switched off, which means that the final Hamiltonian is again given by (4), R and S (after some relaxation time) return to the initial states (7, 12), under influence of the two thermal baths as introduced before. This completes the two-step cycle; now the same heat to work transformation can be repeated.

During the relaxation the thermal baths at temperatures T_h and T_c get, respectively, the amounts of heat

$$Q_h = \text{tr}(H_R \rho) - \text{tr}(H_R \otimes 1 \Omega_{\text{fin}}), \tag{19}$$

$$Q_c = \text{tr}(H_S \sigma) - \text{tr}(H_S \otimes 1 \Omega_{\text{fin}}). \tag{20}$$

The work (18) is expressed as $W_{nm}(\varepsilon, \mu, V) = Q_h + Q_c$.

Following to [10, 23] we determine how Q_h and Q_c relate to the temperatures. Recall that the relative entropy between two density matrices ρ_1 and ρ_2 is defined as [24]

$$S[\rho_1 || \rho_2] = \text{tr}[\rho_1 \ln \rho_1 - \rho_1 \ln \rho_2]. \tag{21}$$

This non-negative quantity reflects the difference between ρ_1 and ρ_2: $S[\rho_1 || \rho_2] = 0$ if and only if $\rho_1 = \rho_2$ [24]. Due to the unitarity of the work-extracting process (17), $\text{tr}[\Omega_{\text{in}} \ln \Omega_{\text{in}}] = \text{tr}[\Omega_{\text{fin}} \ln \Omega_{\text{fin}}]$, where Ω_{in} and Ω_{fin} are defined by (12) and (17). We get

$$-S[\Omega_{\text{fin}} || \Omega_{\text{in}}] = \frac{Q_h}{T_h} + \frac{Q_c}{T_c} \leq 0. \tag{22}$$

If a non-negative amount of work is extracted, $W_{nm}(\varepsilon, \mu, V) = Q_h + Q_c \geq 0$, Eq. (22) implies $W_{nm}(\varepsilon, \mu, V) \leq (1 - \frac{T_c}{T_h})Q_h$. Together with $T_h > T_c$ this means that the heat flows from the higher temperature to the lower one:

$$Q_h \geq 0 \quad \text{and} \quad Q_c \leq 0. \tag{23}$$

For the efficiency of the work-extraction we get that it is always bound from above by the Carnot value

$$\eta_{nm}(\varepsilon, \mu, V) \equiv \frac{W_{nm}(\varepsilon, \mu, V)}{Q_h} \tag{24}$$

$$\leq \eta_{\text{Carnot}} \equiv 1 - \frac{T_c}{T_h}. \tag{25}$$

Eqs.(23–25) are, of course, well known in the ordinary thermodynamics. The advantage of the above derivation is that their validity is confirmed out of the local equilibrium [10] and without requiring a weak system-bath coupling. The latter assumption for deriving the Carnot bound is employed in Ref. [26].

C. Power of work extraction and cycle time

As discussed above, the engine operation consists of two steps: work extraction during a reversible unitary transformation, which (in principle) can be implemented

in an arbitrary short time, and the heat gathering step, when the subsystems R and S relax back to their initial states under influence of the baths. This relaxation is in general irreversible, since, as will be seen below, at the moment when the second stage starts, the state of R (S) as given by (32) is not at local equilibrium with the hot (cold) bath.

If the couplings of R and S to their baths were weak, the second step would take a very long time and thus the power—the work extracted per cycle divided over the cycle time—would tend to zero.

However, the second step can also be implemented in a finite time: We may then assume that the couplings to the bath are finite, or even large, so that the relaxation to the states (7) does not take very long, but the work-extraction stage is still, to a good approximation, thermally isolated. In Appendix A we describe a realistic relaxation mechanism that achieves a finite-time relaxation, but still allows a thermally-isolated work-extracting transformation, as described in section II B 1.

Thus provided we extract a finite amount of work during the first stage, we can get a finite power of work-extraction if the couplings of R and S to their baths are not weak.

III. MAXIMIZING THE WORK OVER THE POTENTIAL $V(t)$

We first optimize the work $W_{nm}(\varepsilon, \mu, V)$ in (18) over all unitary evolutions (17). This is equivalent to maximizing $W_{nm}(\varepsilon, \mu, V)$ over all potentials $V(t)$, defined in (15, 16):

$$W_{nm}(\varepsilon, \mu) \equiv \max_V \left[W_{nm}(\varepsilon, \mu, V) \right] \qquad (26)$$

This is a known problem in thermodynamics [1, 2]. However, the standard textbook answer to this problem applies only to macroscopic, weakly-non-equilibrium systems S and R. The general solution to the maximal work-extraction problem for a given initial state Ω_{in} was given in [25]. For the considered situation (7–12) this solution amounts to the following [25]. Once the initial state (and thus the initial energy) is fixed, for maximizing the work (18) one has to minimize the final energy (cf. (12–14)):

$$\text{tr}[H \Omega_{\text{fin}}] = \sum_{\alpha=1}^{nm} \sum_{\beta=1}^{nm} C_{\alpha\beta} P_\beta E_\alpha, \qquad (27)$$

where

$$C_{\alpha\beta} = \langle E_\alpha | \mathcal{U} | E_\beta \rangle \langle E_\beta | \mathcal{U}^\dagger | E_\alpha \rangle, \qquad (28)$$

Note that $C_{\alpha\beta}$ is a double-stochastic matrix, i.e., *i)* $C_{\alpha\beta} \geq 0$, *ii)* $\sum_{\alpha=1}^n C_{\alpha\beta} = 1$ and $\sum_{\beta=1}^n C_{\alpha\beta} = 1$. Conversely, any matrix that satisfies the features *i)* and *ii)* can be presented as in (28) for some unitary operator $\mathcal{U}$ [34]. A particular case of double-stochastic matrices are permutation matrices, which on each row and on each column contain one element equal to 1, while other elements are equal to zero. For this simplest case, which will be most relevant for our purpose, the relation with unitaries is especially clear, since the transpose Π^T of a permutation matrix Π coincides with its inverse Π^{-1}. Thus, Π is already orthogonal (unitary).

Now it is obvious [34] which double-stochastic matrix $C_{\alpha\beta}$ minimizes the sum in the RHS of (27). This is a permutation matrix that enforces the largest element in E to appear in (27) paired with the smallest element in P. The one but the largest in E gets the one but the smallest element in P, and so on. Finally, the smallest element in P is paired with the largest element in E [34].

To write this down, let Π_H be the permutation that orders the sequence E in (5) in the non-increasing order

$$(\Pi_H E)_{nm} \geq (\Pi_H E)_{nm-1} \geq ... \geq (\Pi_H E)_1, \qquad (29)$$

where $(\Pi_H E)_{nm}$ is the largest element of $\Pi_H E$ which is obviously equal to $\varepsilon_n + \mu_n$ due to (8), $(\Pi_H E)_{nm-1}$ is the one but the largest element of E which in general is not anymore equal to $\varepsilon_n + \mu_{n-1}$, and *etc.*

In the same fashion, let Π_Ω be the permutation that orders in the non-decreasing order the sequence P in (13):

$$(\Pi_\Omega P)_{nm} \leq (\Pi_\Omega P)_{nm-1} \leq ... \leq (\Pi_\Omega P)_1. \qquad (30)$$

Note that $(\Pi_\Omega P)_{nm} = r_n s_m$ and $(\Pi_\Omega P)_{11} = r_1 s_1$ due to (9). Then the minimal final energy can be written as $\sum_{\alpha=1}^{nm} (\Pi_\Omega P)_\alpha (\Pi_H E)_\alpha = \sum_{\alpha=1}^{nm} (\Pi_H^{-1} \Pi_\Omega P)_\alpha E_\alpha$, where Π_H^{-1} is the inverse permutation to Π_H: $\Pi_H^{-1} \Pi_H = 1$. The relation between $\Pi_H^{-1} \Pi_\Omega P$ and E is that if for any two indices $E_\alpha \geq E_\beta$, then $(\Pi_H^{-1} \Pi_\Omega P)_\alpha \leq (\Pi_H^{-1} \Pi_\Omega P)_\beta$.

Thus the maximum of the work W over all interactions $V(t)$ in (15) is equal to [recall (14)]

$$W_{nm}(\varepsilon, \mu) = \sum_{\alpha=1}^{nm} P_\alpha E_\alpha - \sum_{\alpha=1}^{nm} (\Pi_H^{-1} \Pi_\Omega P)_\alpha E_\alpha. \qquad (31)$$

The maximum (31) is achieved for the final state

$$\Omega_{\text{fin}} = \text{diag}[\Pi_H^{-1} \Pi_\Omega P], \qquad (32)$$

which obviously has the same (but differently ordered) eigenvalues as Ω_{in}, since eigenvalues (but not their order) are invariants of any unitary transformation. Thus $\text{tr}(\text{diag}[\Pi_H E] \text{diag}[\Pi_\Omega P]) = \sum_{\alpha=1}^{nm} (\Pi_H^{-1} \Pi_\Omega P)_\alpha E_\alpha$ is the lowest possible energy that can be achieved by permuting the eigenvalues of Ω_{in} [25].

Introducing the separate final states of R and S

$$\rho_{\text{fin}} = \text{tr}_S \Omega_{\text{fin}}, \quad \sigma_{\text{fin}} = \text{tr}_R \Omega_{\text{fin}}, \qquad (33)$$

and using (16) one can write

$$W_{nm}(\varepsilon, \mu) = \text{tr}\left[(\rho - \rho_{\text{fin}})H_R\right] + \text{tr}\left[(\sigma - \sigma_{\text{fin}})H_S\right]. \qquad (34)$$

Note that ρ_{fin} and σ_{fin} commute with the respective Hamiltonians:

$$[\rho_{\text{fin}}, H_R] = 0, \qquad [\sigma_{\text{fin}}, H_S] = 0, \qquad (35)$$

and that they provide a larger probability for a smaller energy, i.e., the analog of (8, 9) is valid for ρ_{fin} and σ_{fin}. Would not these both properties have to hold, one could extract more work from R and S. Note as well that in general ρ_{fin} and σ_{fin} do not have the Gibbsian form, i.e., they are not described by definite temperatures.

The general form of the maximal work extracting interaction $V(t)$ (see (15) for the definition) is given in [25].

IV. MAXIMIZING THE WORK OVER THE SUBSYSTEM SPECTRA

A. Optimal spectral form

We saw how to maximize the work over all interactions $V(t)$. We now maximize the work $W_{nm}(\varepsilon, \mu)$ also over all energy level spacings $\{\varepsilon_k\}_{k=1}^n$ and $\{\mu_l\}_{l=1}^m$, i.e. over the initial state (12). The obtained value of work W will still be a function of T_h, T_c, n and m.

So far we have not been able to carry out analytically the optimization of W over all possible energy spacings. Thus we have to settle for a numerical optimization of W employing the standard optimization packages of Mathematica 5 [1]. The result that emerged for

$$n = m \tag{36}$$

has the following simple form: The maximal work

$$W_{nm} \equiv \max_{\varepsilon, \mu} [W_{nm}(\varepsilon, \mu)] \tag{37}$$

is attained for

$$\varepsilon_n = \varepsilon_{n-1} = ... = \varepsilon_2 > \varepsilon_1 = 0, \tag{38}$$
$$\mu_n = \mu_{n-1} = ... = \mu_2 > \mu_1 = 0, \tag{39}$$

that is when both upper levels ε_2 and μ_2 are $(n-1)$ fold degenerate [2]. This result has been checked numerically for $n = 3, 4, 5$ and we expect it to hold for an arbitrary $n = m$. The intuition behind (38, 39) is that the energy is concentrated—as much as the thermal equilibrium allows—at higher energy levels creating a sort of instability that facilitates the subsequent work-extraction.

For the occupation numbers one gets from (38, 39):

$$r_n = r_{n-1} = ... = r_2 < r_1, \tag{40}$$
$$s_n = s_{n-1} = ... = s_2 < s_1. \tag{41}$$

In the context of (38–41) we note that any state with occupation (40) and energies (38) admits a well defined temperature. While we shall elaborate on this point in

sections IV B 2 and IV B 3 below, it is important to stress already now that the existence of final temperatures came out as the result of maximization: in general the final state (33) do not admit any well-defined temperature.

Once the fact of (38, 39) is recognized, it is not difficult to get the explicit expressions for the maximal work and the corresponding efficiency at maximal work. The maximization of work will be split into two parts. First we shall explore several consequences of conditions (38, 39) treating ε_n and μ_n as parameters. Then in section IV C we shall optimize the work over ε_n and μ_n.

Assuming (38, 39) we get from (4, 38, 39):

$$E = [\underbrace{\underbrace{\varepsilon_2 + \mu_2}_{n-1 \text{ times}}, \varepsilon_2, \underbrace{\mu_2}_{n-1 \text{ times}}, 0]}_{n-1 \text{ times}} \tag{42}$$

Analogously we get from (12, 40, 41):

$$P = [\underbrace{\underbrace{r_2 s_2}_{n-1 \text{ times}}, r_2 s_1, \underbrace{r_1 s_2}_{n-1 \text{ times}}, r_1 s_1]}_{n-1 \text{ times}} \tag{43}$$

When comparing E with P, we see that the action of $\Pi_E^{-1}\Pi_\Omega$, as defined in (29, 30), amounts to interchanging the underlined elements of P in (43):

$$\Pi_H^{-1}\Pi_\Omega P = [\underbrace{\underbrace{r_2 s_2}_{n-1 \text{ times}}, r_1 s_2, \underbrace{r_2 s_1}_{n-1 \text{ times}}, r_1 s_1]}_{n-1 \text{ times}} \tag{44}$$

One gets from (31):

$$W_{nn}(\varepsilon_2, \mu_2) = (n-1)(\varepsilon_2 - \mu_2)(r_2 s_1 - r_1 s_2). \tag{45}$$

This equation can be re-written in a more convenient form by introducing new variables (Boltzmann weights):

$$u \equiv e^{-\beta_h \varepsilon_2}, \qquad v \equiv e^{-\beta_c \mu_2}, \tag{46}$$

so that

$$r_2 = \frac{u}{1 + (n-1)u}, \qquad r_1 = \frac{1}{1 + (n-1)u}, \tag{47}$$

$$s_2 = \frac{v}{1 + (n-1)v}, \qquad s_1 = \frac{1}{1 + (n-1)v}, \tag{48}$$

$$\varepsilon_2 = T_h \ln \frac{1}{u}, \qquad \mu_2 = T_c \ln \frac{1}{v}. \tag{49}$$

Eq. (45) thus reads

$$W_{nn}(u, v) = \frac{(n-1)[u - v]\left[T_h \ln \frac{1}{u} - T_c \ln \frac{1}{v}\right]}{(1 + (n-1)u)(1 + (n-1)v)}. \tag{50}$$

Recalling (6), for the positivity of $W_{nn}(u, v)$ in (50) it is necessary to have [3]

$$u > v \qquad \text{and} \qquad T_h \ln \frac{1}{u} > T_c \ln \frac{1}{v}, \tag{51}$$

---

[1] More specifically, we used three packages based respectively on genetic algorithms, random-gradient search and stimulated annealing, to ensure that we get the correct results.

[2] Appendix B discusses in which sense the system with spectrum (38) is equivalent to a two-level system

[3] The inverse conditions $u < v$ and $T_h \ln \frac{1}{u} < T_c \ln \frac{1}{v}$ are not compatible with each other due to condition (6).

which together with (6, 46, 38, 39) leads to

$$1 < \frac{\varepsilon_2}{\mu_2} < \frac{T_h}{T_c}. \tag{52}$$

For the efficiency of work-extraction we get from (19, 24) and (51): [4]

$$\eta_{nn}(u,v) = 1 - \frac{\min\left[T_h \ln \frac{1}{u}, T_c \ln \frac{1}{v}\right]}{\max\left[T_h \ln \frac{1}{u}, T_c \ln \frac{1}{v}\right]} = 1 - \frac{T_c \ln \frac{1}{v}}{T_h \ln \frac{1}{u}}. \tag{53}$$

1. Carnot limit.

Comparing (50) with (53) we see that for a finite n the efficiency tending to the Carnot value $1 - T_c/T_h$ means that the work extracted per cycle goes to zero [recall that the cycle duration is finite]. In section IV D 2 we show that for $n \to \infty$ this conclusion changes, and a finite amount of work can be extracted with the efficiency [almost] equal to the Carnot value.

Note that the work turning to zero at the Ganot limit is typical for quantum realizations of the Otto cycle [11, 17]. However, there also the cycle duration is very long.

B. Structure of the target state.

1. Optimal work-extracting transformation.

It is seen from (32, 43, 44) that under conditions (38, 39) the optimal work extracting evolution amounts to exchanging the states of S and R:

$$\Omega_{\text{in}} = \rho \otimes \sigma, \qquad \Omega_{\text{fin}} = \sigma \otimes \rho. \tag{54}$$

The meaning of this factorization is discussed after (96).

The optimal work-extraction unitary operation is the so called SWAP, well known—especially for the particular case $n = 2$—as one of the basic gates of quantum computation [24]. Although SWAP is normally realized via composition of several more elementary unitary operations [24], its direct implementation for atomic few-level systems was argued to be feasible [33].

2. Temperatures after work-extraction.

Recall with (6) that Ω_{in} is a non-equilibrium state. The final density matrix Ω_{fin} is also non-equilibrium, but it is expected to be closer to equilibrium than Ω_{in}.

---

[4] Note that from (46) and (53) that for a fixed ε_2 and μ_2 we get the well known Otto-cycle result $\eta = 1 - \frac{\mu_2}{\epsilon_2}$, which is constrained by (52), but otherwise does not depend on temperatures. The Otto cycle is realized via two isothermal and two adiabatically [slow] processes; see, e.g., [11, 17]. Thus in this scenario the power of work is very small. In contrast, as we stressed already, in our situation the efficiency (53) is obtained with a finite power.

Let us see in which sense this expectation is going to be correct. Recalling (10, 11) and (38, 39), the final temperatures T_h' and T_c' of R and S, respectively, is deduced from (54)

$$e^{-\beta_h' \varepsilon_2} = \frac{s_2}{s_1} = v, \quad e^{-\beta_c' \mu_2} = \frac{r_2}{r_1} = u, \tag{55}$$

$$T_h' = T_h \frac{\ln \frac{1}{u}}{\ln \frac{1}{v}}, \quad T_c' = T_c \frac{\ln \frac{1}{v}}{\ln \frac{1}{u}}. \tag{56}$$

With help of (53) and of the Carnot bound (25) we get

$$T_h > T_h', \qquad T_c < T_c'. \tag{57}$$

Thus the initially hotter system R cools down, while initially cooler S heats up. Recall that the physical meaning of the Carnot bound is closely tied to the second law, i.e., to the impossibility of transferring heat completely into work. Eqs. (57) provide a somewhat different perspective on the Carnot bound related to the zeroth law.

Another inequality is derived via (56, 52):

$$\frac{T_h'}{T_c'} < \frac{\varepsilon_2}{\mu_2} < \frac{T_h}{T_c} \frac{T_h'}{T_c'}. \tag{58}$$

This means that the analog of the condition (52) is not satisfied for the final temperatures T_c' and T_h', so that one cannot, again, employ the final state in (54) for work-extraction via a thermally isolated process.

3. Relation with the Curzon-Ahlborn efficiency.

If R initially having a higher temperature than S still has this property after the work extraction, i.e., if

$$T_h' > T_c', \tag{59}$$

we get, employing (56) that (59) is equivalent to

$$\eta_{nn} > 1 - \sqrt{\frac{T_c}{T_h}} = \eta_{\text{CA}}. \tag{60}$$

Here η_{CA} is known as the Curzon-Ahlborn efficiency. For u and v satisfying to the work-extraction condition (51), inequality (60) may or may not be satisfied.

However, below in (65) we show that if the maximal work (over u, v) is extracted, (60) always holds.

Note that according to the Clausius formulation of the second law, if there is a heat exchange between two thermal systems, the heat goes from the hotter system to the colder one. Thus the temperature of the initially hotter system is always larger or equal than the one of the initially colder system. Now recall that the Clausius formulation in the present setup always holds globally in the sense that after the engine cycle is completed, the hotter thermal bath (attached to R) always has lost heat, while the colder bath (attached to S) always has gained heat; see (22, 23). Condition (59) tells that the heat goes from
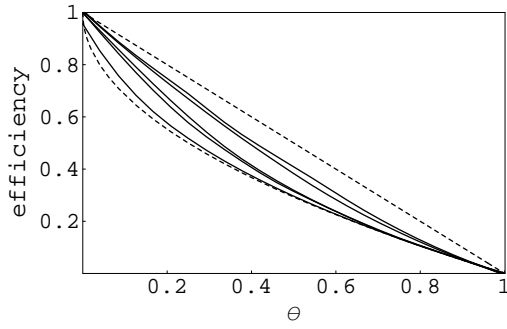
FIG. 2: Normal curves: efficiency η_{nn} at the maximal work versus the temperature ratio $\theta = \frac{T_c}{T_h}$ for various values of the energy level number n. From bottom to top: $n = 2, 501, 1001, 10001, 100001$. Upper dashed line: Carnot efficiency $\eta_{\text{Carnot}} = 1 - \theta$. Lower dashed line: Curzon-Ahlbron efficiency $\eta_{\text{CA}} = 1 - \sqrt{\theta}$. It is seen that $\eta_{\text{Carnot}} > \eta_{nn} > \eta_{\text{CA}}$.

hot to cold not only globally (i.e., for the overall cyclic process of the engine functioning), but also locally, in the work-extraction stage. Indeed, the inverse condition $T_h' < T_c'$ will mean that the initially colder system got hotter at the end of the work-extracting stage. Thus according to (60) the maximal work extraction is related to the local version of the Clausius formulation.

Recall from (52) that for the work extraction at fixed spectra we need a finite difference between the temperatures. Condition (59), which is related to the maximization of work, shows that the final temperatures T_h' and T_c' have also to be different, in contrast to the classic thermodynamic case, where after the maximal work-extraction the overall system has one single temperature [1, 2].

C. Maximization of work over the parameters u and v and efficiency at maximal work.

It remains to maximize the work $W_{nn}(u, v)$ over the parameters u and v:

$$W_{nn} = \max_{u,v}[W_{nn}(u, v)]. \tag{61}$$

Differentiating $W_{nn}(u, v)$ we get the following equations for u_{m} and v_{m}:

$$\frac{1 + (n-1)u_{\text{m}}}{1 + (n-1)v_{\text{m}}} = \sqrt{\theta}\sqrt{\frac{u_{\text{m}}}{v_{\text{m}}}}, \tag{62}$$

$$(1 - \frac{v_{\text{m}}}{u_{\text{m}}})\frac{1 + (n-1)u_{\text{m}}}{1 + (n-1)v_{\text{m}}} = \ln\frac{1}{u_{\text{m}}} - \theta\ln\frac{1}{v_{\text{m}}}, \tag{63}$$

where

$$\theta \equiv \frac{T_c}{T_h} \leq 1. \tag{64}$$

We see that u_{m} and v_{m}—and thus $\frac{W_{nn}(u_{\text{m}}, v_{\text{m}})}{T_h}$ and the efficiency $\eta_{nn}(u_{\text{m}}, v_{\text{m}})$ at the maximal work—depend only on the ratio θ and the number n of the energy levels.
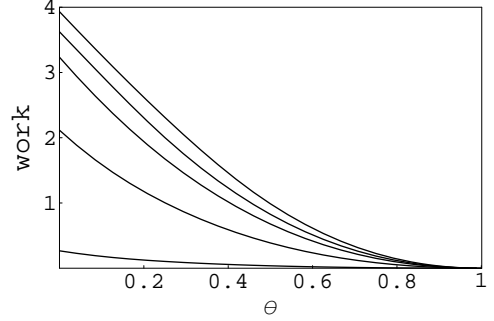

FIG. 3: The dimensionless maximal work $\frac{W_{nn}}{T_h}$ versus the temperature ratio $\theta = \frac{T_c}{T_h}$ for various values of the energy level number n. From bottom to top: $n = 2, 51, 251, 501, 1001$.
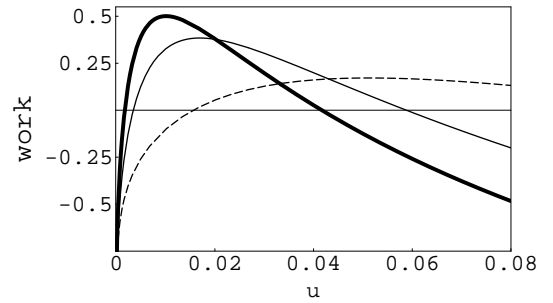

FIG. 4: The dimensionless work $\frac{W_{nn}(u, v_{\text{m}})}{T_h}$ versus the parameter u. The parameter v is set at its optimal value v_{m} that provides the unconditional maximum of $\frac{W_{nn}}{T_h}$. Here $T_c/T_h = 0.5$. Thick curve: $n = 100$. Normal curve: $n = 50$. Dashed curve: $n = 10$. It is seen that the maximum of $\frac{W_{nn}}{T_h}$ gets larger and sharper for larger values of n. The same qualitative conclusion is kept for $\frac{W_{nn}(u_{\text{m}}, v)}{T_h}$.

For $\theta = 1$, Eqs. (62, 63) lead to $u_{\text{m}} = v_{\text{m}}$, and thus to $W_{nn} = \eta_{nn} = 0$, as expected. For a fixed spectrum the efficiency η would not depend on θ at all. This changes upon maximization of W: The behavior of η_{nn} as a function of $\theta = T_c/T_h$ is shown in Fig. 2. While the Carnot efficiency $\eta_{\text{Carnot}} = 1 - \theta$ is always an upper bound for η_{nn}, it appears that η also has a definite lower bound given by the Curzon-Ahlborn efficiency:

$$\eta_{nn} \equiv \eta_{nn}(u_{\text{m}}, v_{\text{m}}) > \eta_{\text{CA}} \equiv 1 - \sqrt{\theta}. \tag{65}$$

In particular, η_{nn} converges toward η_{CA} for $T_h \to T_c$; see next section and Fig. 2. For small values of n, η_{nn} is closer to η_{CA} than to η_{Carnot}. For larger values of n and for θ not very close to 1, η_{nn} is closer to η_{Carnot}.

In addition, Fig. 2 shows that η_{nn} (for a fixed θ) monotonically increases with the number n of energy levels. A similar monotonous increase holds for the dimensionless maximal work W_{nn}/T_h; see Fig. 3.

Fig. 4 presents a projection of the landscape of the dimensionless work $\frac{W_{nn}(u,v)}{T_h}$. We see that upon increasing n, the maximum of $\frac{W_{nn}}{T_h}$ not only gets larger, but it also becomes sharper. (For increasing the visibility of this effect, $\frac{W_{nn}}{T_h}$ is presented as a function of one parameter u, with another being fixed at its optimal value; other reasonable ways of taking projections around the maximum lead to similar qualitative conclusions.) Imagine that various engines with different characteristics u, v and n operate between the two thermal baths. Now the engine with the best characteristics—i.e, larger n and with u and v closer to their optimal values for the given temperatures—will produce the largest amount of work, and will thus over-dominate the others.

D. Special parameter windows

1. Equilibrium limit

For the nearly equilibrium situation $T_h \approx T_c$ one introduces the small parameter ε in the following way

$$\theta \equiv \frac{T_c}{T_h} = 1 - \epsilon, \qquad \frac{u_{\mathrm{m}}}{v_{\mathrm{m}}} = 1 + \epsilon\omega, \qquad (66)$$

Using (66) in (62, 63) and keeping terms $\sim \mathcal{O}(\epsilon)$, we get

$$v_{\mathrm{m}} = e^{-2\omega}, \qquad (67)$$

$$\frac{1}{\omega} = \frac{1 - (n-1)e^{-2\omega}}{1 + (n-1)e^{-2\omega}}. \qquad (68)$$

Eq. (67) expresses v_{m} via ω, while (68) is a transcendental equation for ω. As graphical construction shows, (68) has a unique positive solution for $n \geq 2$. Combining (67, 68) with (50, 53) we get for the maximal work W_{nn} and for the efficiency η_{nn} at the maximal work:

$$W_{nn} = \frac{\epsilon^2\omega^2 e^{-2\omega}}{(1 + ne^{-2\omega})^2} + \mathcal{O}(\epsilon^2), \qquad (69)$$

$$\eta_{nn} = \frac{\epsilon}{2} + \mathcal{O}(\epsilon^2), \qquad (70)$$

where $\omega > 0$ is the solution of (68). As (70) shows, in the considered order η_{nn} coincides with the Curzon-Ahlborn efficiency: $\eta_{\mathrm{CA}} = 1 - \sqrt{1 - \epsilon} = \epsilon/2$.

2. Macroscopic limit: work and efficiency.

If S and R are macroscopic, then $\ln n$ is large:

$$\ln n \gg 1. \qquad (71)$$

Indeed for a system containing $N \gg 1$ particles, the number of energy levels scales as $e^{\mathrm{const}\, N}$.

TABLE I: The maximal work W_{23} and the relative difference $\frac{\eta_{23} - \eta_{\mathrm{CA}}}{\eta_{\mathrm{CA}}}$ between the efficiency at the maximal work and the Curzon-Ahlborn efficiency for various temperatures. Here as in (81), the hot system R has two energy levels, while the cold system S has three energy levels.

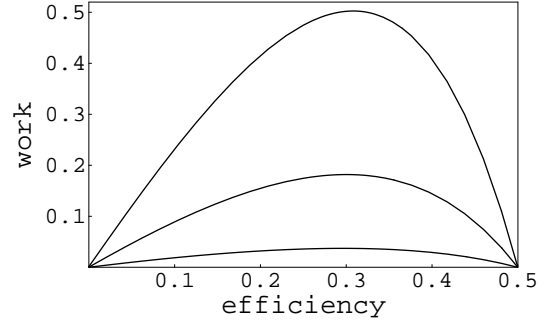|		W_{23}	$\frac{\eta_{23} - \eta_{\mathrm{CA}}}{\eta_{\mathrm{CA}}}$
$T_h = 2$	$T_c = 1$	0.0891062	0.0292807
$T_h = 3$	$T_c = 1$	0.265928	0.0560355
$T_h = 10$	$T_c = 1$	1.97317	0.112379

FIG. 5: The dimensionless work $\frac{W_{nn}(\eta)}{T_h}$ defined in (82) versus the efficiency η. Recall from (82) that $W_{nn}(\eta)$ is obtained by maximizing the work over all parameters for a fixed efficiency η. For this figure $\theta = \frac{T_c}{T_h} = 0.5$, and from bottom to top: $n = 2, 51, 101$. Note that $\frac{W_{nn}(\eta)}{T_h}$ turns to zero two times: for $\eta = 0$ and for $\eta = \eta_{\mathrm{Carnot}} = 1 - \theta$. The maximum of $\frac{W_{nn}(\eta)}{T_h}$ corresponds to the overall maximization of work; see Fig. 3.

Solution of Eqs. (62, 63) is found via expanding over the large parameter $\ln[n-1]$:

$$u_{\mathrm{m}} = \frac{\theta(1-\theta)\ln[n-1]}{n-1} - \frac{\phi + 1 - \theta^2}{n-1}, \qquad (72)$$

$$v_{\mathrm{m}} = \frac{\theta}{(1-\theta)[n-1]\ln[n-1]} + \frac{\phi - 1 + \theta^2}{(1-\theta)^2(\ln[n-1])^2}, \qquad (73)$$

$$\phi \equiv \theta\ln[\theta(1-\theta)\ln[n-1]] + \theta^2\ln\left[\frac{1-\theta}{\theta}\ln[n-1]\right]. \qquad (74)$$

The first terms in the RHS of (72) and (73), respectively, are the dominant ones, while the second terms are the corrections of the relative order $\mathcal{O}\left(\frac{1}{\ln[n-1]}\right)$.

Substituting (72, 73) into (50, 53) we get for the max-

imal work W_{nn} and for the corresponding efficiency η_{nn}

$$\frac{W_{nn}}{T_h} = (1-\theta)\ln[n-1] - \frac{\phi+1+\theta^2}{\theta} \tag{75}$$

$$+ \mathcal{O}\left(\frac{1}{\ln[n-1]}\right), \tag{76}$$

$$\eta_{nn} = 1 - \theta - \frac{2\theta}{\ln[n-1]}\ln[(1-\theta)\ln[n-1]] \tag{77}$$

$$+ \mathcal{O}\left(\frac{1}{(\ln[n-1])^2}\right), \tag{78}$$

where ϕ is given by (74), while the factors $\mathcal{O}\left(\frac{1}{\ln[n-1]}\right)$ and $\mathcal{O}(\frac{1}{(\ln[n-1])^2})$ in (76) and (78) can be recovered with help of the correction in (72) and (73), respectively.

It is seen from (75, 77) that the efficiency converges to the Carnot value, while the maximal work converges to the simple expression

$$W_{nn} = T_h(1-\theta)\ln n, \tag{79}$$

which is a characteristic input thermal energy $T_h \ln n$ times the Carnot efficiency $1-\theta$. The corrections to these results—i.e., the last terms in the RHS of (75, 77)—are important as well. They show that besides (71) one should satisfy

$$(1-\theta)\ln n \gg 1, \tag{80}$$

and that due to negativity of the corrections the actual values of W_{nn} and η_{nn} increase in the macroscopic limit.

3. Macroscopic limit: structure of the engine.

An additional implications of the macroscopic limit is that the temperatures T'_h and T'_c defined in (56) converge to T_h and T_c, respectively. This is clear from (77) and from $\frac{T'_c}{T_c} = \frac{T_h}{T'_h} = \frac{1-\eta_{nn}}{\theta}$, which follows from (56).

As follows from (80), the macroscopic limit is not compatible with the equilibrium limit $T_h \to T_c$. An additional obstacle for the macroscopic limit of the maximal work (but not of the efficiency) is $\theta \to 0$, since then the correction $\frac{1}{\theta}$ in the LHS of (75) diverges.

For $\theta \to 1$ (or $T_c \to T_h$) both the work and efficiency go to zero. However, according to (80) they tend to zero after having left the macroscopic regime (75–78). For this regime (80) implies that the relative difference $1-\theta$ should be smaller than $\frac{1}{\ln n}$, where $\ln n$ has the same order of magnitude as the number of particles.

Note from (72, 73) and from (49) that the energy gaps ε_2 and μ_2 scale as $\ln n$, as should be for a macroscopic systems. Another aspect is uncovered when comparing (72, 73) with (48, 49): in the limit (71) for the high-temperature system R the population is concentrated on the higher ($n-1$ times degenerate) energy level ε_2, while for the low-temperature system S the population is concentrated on the lowest ($\mu_1 = 0$) energy level. In particular, this means that for both S and R in the limit

(71), the canonical ensemble is equivalent (in the leading order of $\frac{1}{\ln n}$) to the micro-canonic ensemble, where all the population is put on one [possibly (quasi)degenerate] energy level. This is what one expects for a macroscopic system. For the spectrum (38, 39), the equivalence of the canonic and micro-canonic ensembles holds under special conditions only, and those conditions are realized after the maximization of the work, as shown by (72, 73).

4. System with $n \neq m$.

So far we studied the case when both subsystems S and R have the same size, i.e. the same number $n = m$ of energy levels. One may want to know what happens for $n \neq m$. One of the simplest models of this type is

$$n = 2 \quad \text{and} \quad m = 3. \tag{81}$$

While for $m = n$ we had the conditions (38, 39), which considerably simplified the subsequent analysis, no analog of these conditions holds for (81). Thus we have to rely on numerical investigations. It appears that we can keep all qualitative conclusions; see Table I.

i) The dimensionless maximal work $\frac{W_{nm}}{T_h}$ is still a function of the temperature ratio $\theta = T_c/T_h < 1$. *ii)* The optimal values of the energy spacings—though not satisfying any simple condition similar to (38, 39)—still depend only on θ. *iii)* The efficiency at the maximal work η_{nm} is still bounded from below by the Curzon-Ahlborn value $1 - \sqrt{\theta}$ and from above by the Carnot value $1 - \theta$.

E. Conditional Maximum

When seeking the maximum of W over the spectra one may impose additional constraints. We study two such scenarios below.

1. Maximization of work for fixed efficiency.

Consider the maximum of $W_{nn}(u, v)$ in (50), when the maximization over the parameters u and v is carried out under the condition of a given efficiency (53). This defines the dimensionless maximal work $\frac{W_{nn}(\eta)}{T_h}$ for a fixed efficiency η:

$$\frac{W_{nn}(\eta)}{T_h} = \max_u\left[\frac{(n-1)\eta[u-u^{\frac{1-\eta}{\theta}}]\ln\frac{1}{u}}{(1+(n-1)u)(1+(n-1)u^{\frac{1-\eta}{\theta}})}\right]. \tag{82}$$

Note that $W_{nn}(\eta)$ is still a function of the temperature ratio $\theta = \frac{T_c}{T_h}$.

As seen in Fig. 5, $\frac{W_{nn}(\eta)}{T_h}$ as a function of η is a bell-shaped curve, which turns to zero once for $\eta = 0$ and then for $\eta = \eta_{\text{Carnot}}$; recall the discussion in section IV A 1.

Fig. 5 shows that $\frac{1}{T_h}W_{nn}(\eta)$ increases upon increasing the number of energy levels, while the maximum η_{nn} of $\frac{1}{T_h}W_{nn}(\eta)$ shifts toward larger values, since in the macroscopic limit $\frac{1}{T_h}W_{nn}(\eta)$ will have its maximum very close to the Carnot value $\eta = 1 - \theta$; recall (77).

2. Emergence of the Curzon-Ahlborn limit during partial optimization of work: An example

It may be instructive to follow in more detail, but via a particular example, how the Curzon-Ahlborn lower bound for the efficiency emerges during a constrained maximization. Assume that the systems R and S have, respectively, the spectra $\{a_R \varepsilon_k\}_{k=1}^n$ and $\{a_S \varepsilon_k\}_{k=1}^n$, which differ only by positive scaling factors a_R and a_S. The initial states of R and S are given by (7). We shall not optimize over the interaction with the work-sources, but rather impose the SWAP transformation (54) for the work-extraction. The work extracted during this transformation is deduced from (18):

$$W_{nn}(a_R, a_S) = (a_R - a_S)\sum_{k=2}^n \varepsilon_k(r_k - s_k), \quad (83)$$

where $r_k = \frac{e^{-\beta_h a_R \varepsilon_k}}{\sum_{l=1}^n e^{-\beta_h a_R \varepsilon_l}}$ and $s_k = \frac{e^{-\beta_h a_S \varepsilon_k}}{\sum_{l=1}^n e^{-\beta_h a_S \varepsilon_l}}$. In the second step the systems R and S are subjected to free relaxation, as described in section II B 2. Without loss of generality we assume $\frac{T_c}{T_h} < \frac{a_S}{a_R}$, which means that R (S) has higher (lower) temperature [5]. The efficiency of work-extraction is obtained from (83, 19, 24):

$$\eta_{nn}(a_R, a_S) = 1 - \frac{a_S}{a_R}. \quad (84)$$

Note that the duration of the work extracting cycle is finite. Instead of doing the full optimization over a_S and a_R, we now employ a necessary conditions $\frac{\partial W_{nn}}{\partial a_R} = \frac{\partial W_{nn}}{\partial a_S} = 0$, for the work (83) to be optimal over these two parameters. Working out $\frac{\partial W_{nn}}{\partial a_R} + \frac{\partial W_{nn}}{\partial a_S} = 0$ we get

$$1 - \frac{a_S}{a_R} = 1 - \sqrt{\frac{T_c C_c}{T_h C_h}}, \quad (85)$$

where $C_h = \beta_h^2 a_R^2 [\langle \varepsilon^2 \rangle_h - \langle \varepsilon \rangle_h^2]$ and $C_c = \beta_c^2 a_S^2 [\langle \varepsilon^2 \rangle_c - \langle \varepsilon \rangle_c^2]$ are the equilibrium specific heats of R and S, respectively, calculated for the initial states at the temperatures T_h and T_c (recall that $T_h > T_c$).

For a finite number of energy levels n, the equilibrium specific heat C is a non-monotonous function of temperature: for low temperatures C naturally goes to zero, while

---

[5] Since we did not optimize over the interaction with the sources of work, $W_{nn}(a_R, a_S)$ in (83) can be negative, e.g., for $T_h = T_c$, which means that the work is put into the system.

---

for for very large temperatures it is zero again, since there is a maximal energy a finite-level system can have. Thus for a finite n, the ratio $\frac{C_c}{C_h}$ may be larger or smaller than one, so that (85) does not imply any definite relation between the actual efficiency (84) and the Curzon-Ahlborn value. However, for $n = \infty$ the equilibrium specific heat can be [but need not be] a monotonically increasing function of the temperature, since now the system can accept infinitely large energies. An example of such a behavior is given by a harmonic oscillator with frequency ω, where $\varepsilon_n = \hbar\omega n$, and the specific heat is $C = b^2 e^b (e^b - 1)^{-2}$, with $b = \hbar\omega/T$. Thus, now $\frac{C_c}{C_h} < 1$, since $T_h > T_c$, and (85) implies that the Curzon-Ahlborn value is a lower bound for the efficiency (84).

We draw two conclusions: *i)* upon partial maximization the Curzon-Ahlborn value may or may not be a lower bound for the efficiency depending on the system details; *ii)* the emergence of the Curzon-Ahlborn lower bound is facilitated for systems with classical (i.e., unlimited from above) spectrum, since now the equilibrium specific heat can be an increasing function of the temperature.

V. OPTIMIZING THE SECOND STEP

Now we want to extract work also in the second stage, i.e., during the relaxation of R and S. In contrast to the previous consideration, here we *impose* that the couplings of R and S with their respective baths are weak. We start at the initial state (12) of S+R. The maximal work-extracting transformation is applied, the work (31) is extracted, and R (S) is left in the final state ρ_{fin} (σ_{fin}), as given by (33). Recall from (35) that ρ_{fin} and σ_{fin} commute with their Hamiltonians. Due to the weak-coupling assumption the overall initial state of R (S) and the hot (cold) bath is, to a good approximation, factorized: $\rho \otimes \rho_B$ ($\sigma \otimes \sigma_B$), where ρ_B (σ_B) is the equilibrium state of the hot (cold) bath. Since the interaction with the bath is not essential during the first [thermally isolated] stage, the same factorization holds after this stage: $\rho_{\text{fin}} \otimes \rho_B$ and $\sigma_{\text{fin}} \otimes \sigma_B$.

If we just let R (S) to interact with the thermal bath at temperature T_h (T_c), it will relax back to the initial state ρ (σ). We can however control this relaxation process and extract an additional amount of work.

According to the standard thermodynamic argument [1, 2], the *maximal* work extractable—via a cyclic-Hamiltonian process—from a system in the initial state ρ_{fin} in contact with the thermal bath at temperature T_h is bounded from above by the free energy difference

$$\Delta F_R = \text{tr}[\rho_{\text{fin}}H_R] + T_h \text{tr}[\rho_{\text{fin}}\ln\rho_{\text{fin}}]$$
$$- \text{tr}[\rho H_R] - T_h \text{tr}[\rho\ln\rho] \equiv \Delta U_R - T_h\Delta S_R, \quad (86)$$
$$\Delta U_R \equiv \text{tr}[\rho_{\text{fin}}H_R] - \text{tr}[\rho H_R], \quad (87)$$
$$\Delta S_R \equiv -\text{tr}[\rho_{\text{fin}}\ln\rho_{\text{fin}}] + \text{tr}[\rho\ln\rho], \quad (88)$$

where ρ is the equilibrium state (7), and where ΔU_R and ΔS_R are the changes of the energy and the von Neumann

TABLE II: The maximal value of $\mathfrak{W}_{33}$, given by (99) with maximization over the parameters u and v, compared to the maximal value of W_{33}; see (61). The maximization is done for $n = m = 3$ (two three-level systems). We also present the relative difference $\frac{\eta_{33} - \eta_{\mathrm{CA}}}{\eta_{\mathrm{CA}}}$ between the efficiency η_{33} at the maximal work W_{33} and the Curzon-Ahlborn efficiency $\eta_{\mathrm{CA}} = 1 - \sqrt{T_c/T_h}$. Recall that the work $\mathfrak{W}_{33}$ is extracted with the Carnot efficiency; see (95, 97).

	W_{33}	$\frac{\eta_{23} - \eta_{\mathrm{CA}}}{\eta_{\mathrm{CA}}}$	$\mathfrak{W}_{33}$
$T_h = 2 \quad T_c = 1.7$	0.009271	0.000931	0.0198908
$T_h = 2 \quad T_c = 1$	0.128486	0.0142578	0.478908
$T_h = 10 \quad T_c = 1$	3.07425	0.0711658	9.20744

entropy, respectively.

A detailed discussion of (88) and its derivation is given in Appendix C 1. Moreover, in Appendix C 2 we show that if the initial state commutes with the Hamiltonian, $[\rho_{\mathrm{fin}}, H_{\mathrm{R}}] = 0$, as is the case according to (35), the bound (88) is achieved via:

(1.) *Suddenly* changing the energy spacings of H_{R} according to (C17, C18, C19). This brings R to the local equilibrium with the hot bath, since at the end of the sudden change the temperature of R is equal to the bath temperature T_c. The change is sudden as compared to the relaxation time induced by the weak coupling to the hot bath, so that the interaction with the hot bath during the change is neglected.

(2.) *Slowly* changing the spacings back to their initial values. Now the change is much slower than the relaxation time of R driven by the weak-coupling with the hot bath. Thus R is always in the local equilibrium with the hot bath; see the discussion around (C21). The equilibrium state ρ is attained at the end of this isothermal process. The work-extraction in this stage takes a time much longer than the relaxation time of R which [due to the weak-coupling assumption] is already larger than internal times of R. Thus the overall process has a small work extraction power.

The heat received from the hot thermal bath is (compare with (19)):

$$Q_h = \mathrm{tr}(H_{\mathrm{R}}\,\rho) - \mathrm{tr}(H_{\mathrm{R}}\rho_{\mathrm{fin}}) + \Delta F_{\mathrm{R}} = -T_h \Delta S_{\mathrm{R}}, \quad (89)$$

where ΔF_{R} is given by (88).

The same procedure is applied to S thereby extracting the work ΔF_{S}:

$$\begin{aligned}
\Delta F_{\mathrm{S}} &= \mathrm{tr}[\sigma_{\mathrm{fin}} H_{\mathrm{S}}] + T_c \mathrm{tr}[\sigma_{\mathrm{fin}} \ln \sigma_{\mathrm{fin}}] \\
&\quad - \mathrm{tr}[\sigma H_{\mathrm{S}}] - T_c \mathrm{tr}[\sigma \ln \sigma] \equiv \Delta U_{\mathrm{S}} - T_c \Delta S_{\mathrm{S}}, \quad (90)
\end{aligned}$$

where ΔU_{S} and ΔS_{S} are defined analogously to (88):

$$\Delta U_{\mathrm{S}} \equiv \mathrm{tr}[\sigma_{\mathrm{fin}} H_{\mathrm{S}}] - \mathrm{tr}[\sigma H_{\mathrm{S}}], \quad (91)$$

$$\Delta S_{\mathrm{S}} \equiv -\mathrm{tr}[\sigma_{\mathrm{fin}} \ln \sigma_{\mathrm{fin}}] + \mathrm{tr}[\sigma \ln \sigma], \quad (92)$$

The total work extracted during the cycle reads

$$\begin{aligned}
\mathfrak{W}_{nm}(\varepsilon, \mu) &= W_{nm}(\varepsilon, \mu) + \Delta F_R + \Delta F_S \quad (93) \\
&= -T_h \Delta S_{\mathrm{R}} - T_c \Delta S_{\mathrm{S}}, \quad (94)
\end{aligned}$$

where $W_{nm}(\varepsilon, \mu)$ is the maximal work (34) [optimized over the interaction with the sources of work] extracted during the first stage. Note that the only feature of $W_{nm}(\varepsilon, \mu)$ employed in passing from (93) to (94) is that the evolution of R+S in the first stage is themally isolated, i.e., $W_{nm}(\varepsilon, \mu)$ is equal to the energy difference of R+S; see (18).

The efficiency that follows from (89, 94) is

$$\mathfrak{e}_{nm}(\varepsilon, \mu) = 1 + \frac{T_c}{T_h} \frac{\Delta S_{\mathrm{S}}}{\Delta S_{\mathrm{R}}}. \quad (95)$$

The work (94) is now optimized over the spectra $\varepsilon_n, ..., \varepsilon_2$ and $\mu_n, ..., \mu_2$ of R and S, respectively, for fixed temperatures.

For R and S having the same number of energy levels: $n = m$, the numerical optimization of (94) produced the same result (38–41) as for the work extraction at a finite power: the optimal work is achieved for the upper levels being $n - 1$ times degenerate.

As we already know, Eqs. (38, 39) imply the factorized final state (32): $\rho_{\mathrm{fin}} = \sigma$ and $\sigma_{\mathrm{fin}} = \rho$; see (54). Recalling the definitions (88, 92) of ΔS_{S} and ΔS_{R}, we get

$$\Delta S_{\mathrm{S}} + \Delta S_{\mathrm{R}} = 0. \quad (96)$$

The physical meaning of this condition becomes clear when noting that if the factorization (54) was invalid—i.e. if there were correlations in the final state (32)—the sum $\Delta S_{\mathrm{S}} + \Delta S_{\mathrm{R}}$ would be larger than zero, as implied by the sub-additivity of the von Neumann entropy[6]. Since after the second step of the engine operation, the systems R and S return to their initial states (7), the surplus entropy $\Delta S_{\mathrm{S}} + \Delta S_{\mathrm{R}}$ has to be consumed by the thermal baths, thereby increasing their entropy. Thus, the factorization condition (54) eliminates this potential channel for entropy generation.

Eqs. (95) and (96) imply that the efficiency is equal to the *maximally possible* Carnot value:

$$\mathfrak{e}_{nn} = \eta_{\mathrm{Carnot}} = 1 - \frac{T_c}{T_h}. \quad (97)$$

Using (46–49) we get for the work (94)

$$\mathfrak{W}_{nn}(u, v) = (T_h - T_c)\left(\mathrm{tr}[\sigma \ln \sigma] - \mathrm{tr}[\rho \ln \rho]\right) \quad (98)$$

$$\begin{aligned}
&= (T_h - T_c)\left[\ln \frac{1 + (n-1)u}{1 + (n-1)v}\right. \\
&\quad \left. + (n-1)\frac{v \ln v - u \ln u + (n-1)uv \ln \frac{v}{u}}{(1 + (n-1)u)(1 + (n-1)v)}\right]. \quad (99)
\end{aligned}$$

The latter expression is to be optimized over u and v. It should be clear from our constructions that

$$\mathfrak{W}_{nn} \equiv \max_{u,v}[\mathfrak{W}_{nn}(u, v)] \geq W_{nn}. \quad (100)$$

---

[6] If $S_{12} = -\mathrm{tr}\rho_{1+2} \ln \rho_{1+2}$ is the von Neumann entropy of a composite system with density matrix ρ_{1+2}, then the sub-additivity implies [24]: $S_{1+2} \leq S_1 + S_2$, where $S_i = -\mathrm{tr}\rho_i \ln \rho_i$, $i = 1, 2$, and where $\rho_1 = \mathrm{tr}_2 \rho_{1+2}$, $\rho_2 = \mathrm{tr}_1 \rho_{1+2}$.

It is seen that $\frac{\mathfrak{W}_{nn}}{T_h}$—analogously to W_{nn} in (61)—is a function of the temperature ratio $\theta \equiv \frac{T_c}{T_h}$.

The difference $\mathfrak{W}_{nn} - W_{nn}$ quantifies how much work has been traded in by sacrificing the power. As Table II shows, especially for $T_h \approx T_c$, the ratio $\frac{\mathfrak{W}_{nn}}{W_{nn}}$ can approach ~ 20.

A. Can one extract a finite amount of work at a finite power and with a finite-level engine?

Recall that in section IV A 1 we discussed that when a finite-level engine approaches the Carnot efficiency at a finite power, the work extracted per cycle vanishes, though the cycle duration is finite. This is not anymore the case for a macroscopic engine, as we saw in section IV D 2. Above we have shown that if one imposes the zero-power condition, the work-extraction at the Carnot efficiency is possible.

Can one extract a finite amount of work at a finite power and with a finite-level engine? In other words, is it possible to extract in the second stage the amounts of work (88, 90) without employing very long processes? The answer is yes [38]. A finite cycle time is achieved, since the coupling to the bath is kept weak before and after the work-extraction process, but it is made finite in the intermediate stages. However, the scheme proposed in [38] has a serious drawback: one should employ the sources of work-extraction that act not only on R and S, but also on the corresponding thermal baths. The latters are normally out of any direct control, and it is not clear to what extent the scheme proposed in [38] is realistic. Thus the question is open.

VI. DISCUSSION AND CONCLUSIONS

A. Summary.

Our purpose was in deducing the physics of quantum heat engines from maximizing the work [extracted per cycle] under various constraints. We have found that the work maximization can be introduced on three different levels:

I. Our model is described in section II; see also Fig. 1. One optimizes the extracted work over the interaction of the quantum systems S and R with the external source of work. In general, this procedure is not sufficient for extracting a finite amount of work at $T_h \neq T_c$. In addition, the intermediate stages of the engine functioning are not described by well defined temperatures: if S and R had well defined temperatures before the work-extraction, they are not guaranteed to have definite temperatures after (or during) this process.

II. One maximizes the work, in addition, over the spectral structure of S and R. This turns out to be a crucial step, since after this optimization one finds that *i)* S and R are described by well defined temperatures in the intermediate stages of the engine operation. *ii)* The Clausius formulation of the second law—heat goes from higher to lower temperature—is satisfied not only for the total work-extracting cycle [as it should], but also locally for the intermediate stages of the process. Thus the local thermodynamic structure emerges as a result of work optimization. *iii)* The efficiency at the maximal work is bounded from *below* by the Curzon-Ahlborn efficiency $1 - \sqrt{T_c/T_h}$. This is in addition to the upper Carnot bound $1 - T_c/T_h$, which holds for this model generally, i.e., with or without any optimization. The important feature of these limits is that they are completely system-independent.

Note that the duration of the engine cycle is finite. Thus the power of work is finite and [within our model] tunable to a large extent; see section II C for details.

III. The Curzon-Ahlborn bound is approached close to equilibrium $T_h \to T_c$, while the Carnot upper bound is reached—at a finite power—for a large system S and R. In the latter macroscopic limit S and R (tuned to extract the maximal work) have the expected features of macroscopic systems, e.g., canonic and micro-canonic ensembles are equivalent for them. The equilibrium and macroscopic limits do not commute.

IV. Finally, one optimizes the extracted work also over the interaction of S and R with their respective thermal baths. This optimization was carried out via increasing the cycle duration and thus reducing the power of work extraction. The main result is that the Carnot value for the efficiency—with a finite amount of work extracted at a small power—is reached for finite systems S and R. Thus there are two options for operating close to the Carnot efficiency: either one extracts a finite amount of work per a very long cycle, or a small amount of work per a finite cycle.

B. Relation to previous work

1. Curzon-Ahlborn efficiency.

The Curzon-Ahlborn efficiency has repeatedly been found as the (approximate) efficiency of model macroscopic heat engines functioning at a finite power [27, 28]. Ref. [7] studies a quantum heat engine model, and derives the Curzon-Ahlborn value, but in the limit of vanishing interaction with sources of work. Taking up the early approach of [27], it was stated recently that the Curzon-Ahlborn value is an *upper* bound for the efficiency of a heat engine operating at the maximum power and according to the rules of linear transport theory [29].

In contrast, we show that the Curzon-Ahlborn value is the *lower* bound for the efficiency at the maximal work and a finite power of work. This clarifies the meaning of the Curzon-Ahlborn value, since our derivation does not employ the linear transport theory, and is based on the fundamental level of quantum mechanics. It is important that rather different physical mechanisms can lead to the

same expression of the Curzon-Ahlborn value.

2. Carnot efficiency.

According to standard thermodynamics the efficiency of any heat engine is bounded by the Carnot value [1, 2]. Although this result was derived within (nearly)equilibrium thermodynamics, it can be extended to a rather general class of non-equilibrium heat engines; see (22, 25) and [10]. While useful as an upper bound, the Carnot efficiency by itself is often considered to lack practical significance, since the the power of work goes to zero in this limit. This has been a motivation to develop the so-called finite-time [finite-power] thermodynamics [30].

The quantum situation can be even more restrictive in this context, since typical implementations are based on the Otto cycle, for which not only the power, but also the extracted work goes to zero when approaching the Carnot bound [11, 17, 20].

For a finite-level quantum heat engine operating close to the Carnot efficiency we have two related results. First we show that for such an engine operating with a finite cycle duration, the work extracted per cycle goes to zero when the efficiency approaches the Carnot value; see our discussion at the end of section IV A. Second result is that the same quantum engine can extract a non-zero amount of work—but with a long cycle duration—sharply at the Carnot efficiency, provided one properly optimizes also over the second step, the bath-engine interaction.

Results on small quantum [5, 12, 13, 15, 16] or classical [31, 32] engines working at zero power, but with the Carnot efficiency were already reported in literature.

Most of these results [5, 15, 16, 31, 32] concern continuously [not cyclically] working engines, where the external source of work is absent and the work is performed by moving particles against an external force. In contrast, Refs. [12, 13] report on a cyclically operating quantum heat engine that [at a small power] deliver with the Carnot efficiency a finite amount of work per cycle to an external source of work. Both these works attempt to model quantum Carnot cycle.

We also show that the macroscopic engine can extract a finite amount of work at a finite-power and with [almost] the Carnot efficiency. To our knowledge this effect was so far never reported in the literature.

C. Open questions.

We leave open two important questions:

(1.) How the engine tunes itself to the maximal work-extraction regime? Some hints on this point were expressed at the end of section IV C.

(2.) Can one extract work at a finite power by employing a finite-level engine operating at the Carnot efficiency? This question was discussed in section V A.

Acknowledgments

[1] L.D. Landau and E.M. Lifshitz, *Statistical Physics, I*, Pergamon Press Oxford, 1978.

[2] H.B. Callen, *Thermodynamics*, (John Wiley, NY, 1985).

[3] A.I. Zotin and I. Lamprecht, J. Theor. Biol. **180**, 207 (1996).
I. Lamprecht and A.I. Zotin, eds., *Thermodynamics and Kinetics of Biological Processes* (Walter de Gruyter, Berlin, 1982).

[4] H. Ozawa *et al.*, Rev. Geophys. **41**, 4 (2003).

[5] H.E.D. Scovil and E.O. Schultz-DuBois, Phys. Rev. Lett., **2**, 262 (1959).
J.E. Geusic et al., Phys. Rev., **156**, 343 (1967).

[6] V.K. Konyukhov and A.M. Prokhorov, Uspekhi Fiz. Nauk, **119**, 541 (1976).

[7] R. Kosloff, J. Chem. Phys. 80, 1625 (1984).

[8] E. Geva and R. Kosloff, J. Chem. Phys. **104**, 7681 (1996).
T. Feldmann and R. Kosloff, Phys. Rev. E **70**, 046110 (2004).

[9] C.M. Bender, D.C. Brody and B.K. Meister, J. Phys. A **33**, 4427 (2000). S. Lloyd, Phys. Rev. A, **56** (1997) 3374. J. He et al., Phys. Rev E, **65**, 036145 (2002).

[10] A.E. Allahverdyan, R. Balian and Th.M. Nieuwenhuizen,

[11] H. T. Quan, Y. D. Wang, Y. Liu, C.P. Sun and F. Nori, Phys. Rev. Lett. **97**, 180402 (2006).

[12] H. T. Quan, Y. D. Wang, Y. Liu, C.P. Sun and F. Nori, quant-ph/0611275.

[13] G.P. Beretta, quant-ph/0703261.

[14] E. Boukobza and D. Tannor, Phys. Rev. A **74**, 063823 (2006); Phys. Rev. A **74**, 063822 (2006).

[15] T.E. Humphrey, R. Newbury, R.P. Taylor and H. Linke, Phys. Rev. Lett. **89**, 116801 (2002).

[16] T.E. Humphrey and H. Linke, Physica E, **29**, 390 (2005).

[17] T. D. Kieu, Phys. Rev. Lett. **93**, 140403 (2004); Eur. Phys. J. D **39**, 115 (2006).

[18] M.O. Scully, M.S. Zubairy, G.S. Agarwal, and H. Walther, Science **299**, 862 (2003).
M. Scully, Phys. Rev. Lett. **87** (2001) 220601; *ibid* **88** (2002) 050602.

[19] A. E. Allahverdyan, R. Serral Gracia, and T. M. Nieuwenhuizen, Phys. Rev. E **71**, 046106 (2005).

[20] M. J. Henrich, M. Michel and G. Mahler, Europhys. Lett., **76**, 1057 (2006).
M. J. Henrich, G. Mahler and M. Michel, cond-

J. Mod. Opt. **51**, 2703 (2004).

mat/0705.0075.

[21] R. Alicki, M. Horodecki, P. Horodecki and R. Horodecki, Open Syst. Inf. Dyn. **11**, 205 (2004).

[22] K. Maruyama, F. Nori and V. Vedral, physics/07073400.

[23] H.M. Partovi, Phys. Lett. A, **137**, 440 (1989).

[24] A. Galindo and M.A. Martin-Delgado, Rev. Mod. Phys., **74**, 347, (2002).

[25] A.E. Allahverdyan, R. Balian and Th.M. Nieuwenhuizen, Europhys. Lett. **67**, 565 (2004).

[26] R. Alicki, J. Phys. A, **12**, L103 (1979).

[27] F. Curzon and B. Ahlborn, Am. J. Phys. **43**, 22 (1975).

[28] H. Leff, Am. J. Phys. **55**, 602 (1987).
P. T. Landsberg and H. Leff, J. Phys. A **22**, 4019 (1989).

[29] C. Van den Broek, Phys. Rev. Lett. **95**, 190602 (2005).

[30] R.S. Berry, V.A. Kazakov, S. Sieniutycz, Z. Szwast, and A.M. Tsvilin, *Thermodynamic Optimization of Finite–Time Processes* (John Wiley & Sons, Chichester, 2000).

[31] H. Sakaguchi, J. Phys. Soc. Jpn. **67**, 709 (1998).

[32] I. Derenyi and R. D. Astumian, Phys. Rev. E **59**, R6219 (1999).
M. Asfaw and M. Bekele, Eur. Phys. J. B **38**, 457, (2004); arXiv:cond-mat/0605233.

[33] N. Sangouard, X. Lacour, S. Guerin, and H.R. Jauslin, quant-ph/0505163.

[34] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, (Academic Press, New York, 1979).

[35] A B Brailovskii, V.L. Vaks and V.V. Mityugov, Physics Uspekhi, **166**, 795 (1996).

[36] V.V. Mityugov, Physics Uspekhi, **170**, 681 (2000).

[37] V. Scarani, M. Ziman, P. Stelmachovic, N. Gisin, and V. Buzek, Phys. Rev. Lett. **88**, 097905 (2002).

[38] A.E. Allahverdyan, unpublished notes.

APPENDIX A: COLLISIONAL RELAXATION.

The purpose of this discussion is to outline a realistic example of a finite-duration [collisional] relaxation process which is consistent with the thermally isolated work extracting setup described in section II B 1. Our presentation follows to [23, 35, 36].

The thermal bath is modeled as a collection of $N \gg 1$ independent equilibrium systems (particles) with density matrices $\omega_i = \frac{1}{Z_i} \exp[-\beta H_i]$ and Hamiltonians H_i, where $i = 1, .., N$, and where $1/\beta = T$ is the bath temperature. This formalizes the intuitive notion of the bath as a collection of many weakly-interacting equilibrium systems.

The target system R starts in [an arbitrary] initial state ρ and has Hamiltonian H_R. The collisional relaxation is realized when the particles of the bath sequentially interact [collide] with R. Consider the first collision. The initial state of R and the first bath particle is $\Omega_{1+R} = \rho \otimes \omega_1$, the interaction between them is realized via a unitary operator $\mathcal{V}$, so that the final state after the first collision is $\Omega'_{1+R} = \mathcal{V}\Omega_{1+R}\mathcal{V}^\dagger$. Define separate final states:

$$\rho' = \mathrm{tr}_1 \Omega'_{1+R}, \qquad \omega'_1 = \mathrm{tr}_R \Omega'_{1+R}, \qquad (A1)$$

where tr_1 and tr_R are the partial over the first particle and R, respectively. Recall the definition (21) of the relative entropy. The unitarity of $\mathcal{V}$ implies

$$
\begin{aligned}
S[\Omega'_{1+R} \| \rho' \otimes \omega_1] &= \mathrm{tr}[\Omega_{1+R} \ln \Omega_{1+R}] \\
&\quad - \mathrm{tr}[\Omega'_{1+R} \ln(\rho' \otimes \omega_1)]. \quad (A2)
\end{aligned}
$$

Employing $\omega_1 = \frac{1}{Z_1} \exp[-\beta H_1]$ and $S[\Omega'_{1+R} \| \rho' \otimes \omega_1] \geq 0$ in (A2) we get

$$T\Delta S_R + \Delta U_1 \geq 0, \qquad (A3)$$

where $\Delta S_R = \mathrm{tr}[-\rho' \ln \rho' + \rho \ln \rho]$ and $\Delta U_1 = \mathrm{tr}(H_1[\omega'_1 - \omega_1])$ are, respectively, the change of the entropy of R and the average energy of the first particle.

We now require that the that interaction $\mathcal{V}$ conserves the energy: $\Delta U_1 = -\Delta U_R$. Using this in (A3) one has

$$\Delta U_R - T\Delta S_R \leq 0. \qquad (A4)$$

Since we did not use any special feature of the initial state of R, (A4) holds for subsequent collisions of R with the bath particles. Thus $\Delta U_R - T\Delta S_R$ decays in time, and it should attain its minimum. It is well-known [1] that this minimum is reached for the Gibbs matrix $\rho \propto e^{-\beta H_R}$: the collisions drive R to equilibrium starting from an arbitrary state. The coupling with the bath is switched on during the collision only, but since this coupling can be sizable, the relaxation time can be very short [23].

Further results and concrete scenarios of collisional relaxation are given in [23, 35–37]. This includes the rate of [exponential] convergence to equilibrium which was favourably compared to experiments in [35].

As for the implementation of the work extracting pulse described in section II B 1, one notes that within the present relaxation model, the duration of the pulse has to be shorther than the time between the collisions. If the latter time is much larger than both the collision time and the pulse time, almost any implementation of the pulse will amount to a thermally isolated process.

APPENDIX B: SYSTEM WITH $n-1$ FOLD DEGENERATE SPECTRUM.

Consider an n-level quantum system R whose upper $n-1$ energy levels are degenerate; see (38). Let us show that the behavior of R is equivalent to that of a two-level system ($n = 2$) with respect to all transformations that do not resolve the $n-1$ fold degeneracy of the spectrum (38). To this end we introduce for $n = 3$ generalized Pauli matrices [generalizing to $n > 3$ is straightforward]

$$\hat{\sigma}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \hat{\sigma}_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \hat{\sigma}_2 = i\hat{\sigma}_1\hat{\sigma}_3. \text{ (B1)}$$

It is clear that *i)* the algebra of these matrices is identical to that of the corresponding Pauli matrices; *ii)* Any perturbation of the Hamiltonian H_R of R by an arbitrary

linear combination of $\{\hat{\sigma}_i\}_{i=1}^3$ does not change the double-degeneracy of H_R; *iii)* If the initial state of R can be expressed only via $\{\hat{\sigma}_i\}_{i=1}^3$, then during any interaction with another quantum system—which is written only via $\hat{\sigma}_i$ and arbitrary operators of this system—R is dynamically equivalent (via the Heisenberg representation) to a two-level system. An example of such an initial state is when R is described by a definite temperature.

APPENDIX C

1. Free energy bound for the maximal work extractable from a system in contact with a bath.

We are given a quantum system in a non-equilibrium initial state Φ and Hamiltonian H. Some work is to be extracted from this system via a cyclic-Hamiltonian,

$$H(0) = H(\tau) = H, \tag{C1}$$

thermally isolated process. We want to give an upper bound for this work. The work-extraction process is unitary, and thus it conserves the eigenvalues of Φ. In particular, it conserves the von Neumann entropy

$$S[\Phi] = -\mathrm{tr}\,[\Phi \ln \Phi]. \tag{C2}$$

Recall that the (positive) extracted work is defined as the difference between the initial and final energies. We can give an upper bound for the maximal extractable work by looking for a final state with the *minimal* energy compatible with the above conservation of the von Neumann entropy [1, 2]. The corresponding minimization procedure is standard in statistical physics [1]—since it is dual to the maximization of the entropy for a fixed average energy—and it produces for the final state a Gibbsian density matrix

$$\Phi(\beta_f) = \frac{1}{Z}\, e^{-\beta_f H}, \qquad Z = \mathrm{tr}\, e^{-\beta_f H}, \tag{C3}$$

with an inverse final temperature β_f defined via the entropy conservation:

$$S[\Phi] = S[\Phi(\beta_f)]. \tag{C4}$$

The upper bound for the maximal work is now

$$W_{\mathrm{th}} = \mathrm{tr}\,[H\Phi] - \mathrm{tr}\,[H\Phi(\beta_f)]. \tag{C5}$$

This is the standard solution of the maximal work extraction problem given in thermodynamics [1, 2].

Let us now adopt the above reasoning assuming that the overall system consists of a large thermal bath B and a small system R. Initially B is in equilibrium at inverse temperature β, while R is in an arbitrary state ρ. Thus the initial state is

$$\Phi = \rho \otimes \rho_B(\beta), \qquad \rho_B(\beta) = \frac{1}{Z_B}\, e^{-\beta H_B}, \tag{C6}$$

where H_B is the bath Hamiltonian. The overall initial Hamiltonian is

$$H = H_R \otimes 1 + 1 \otimes H_B, \tag{C7}$$

where H_R is the Hamiltonian of R.

We now employ the following facts:
(1.) The number of bath degrees of freedom is much larger than that for the small system R.
(2.) Since the bath started from a passive equilibrium state, the extracted work is expected to be of the same order as the energy of R. Thus the extracted work is much smaller than the bath energy.
(3.) Since $\beta = \mathcal{O}(H_R)$, the difference between β and β_f is expected to be much smaller than H_R.
(4.) The interaction between B and R occurs only during the work-extraction process. It is negligible both before and after the process. The interaction may be sizable during the work-extraction, but the corresponding energy costs for switching it on and off are already included in the work.

Let us now write (C4) as

$$S[\rho] + S[\rho_B(\beta)] = S[\rho(\beta)] + S[\rho_B(\beta_f)], \tag{C8}$$

where $\rho(\beta) \propto \exp[-\beta H_R]$, and where the LHS follows from the fact that initially the bath and the small system were in a factorized state. The RHS follows from the Gibbs density matrix (C3) of the final state, taking into account in the final state R and B do not interact.

In (C8) we neglected the difference between β and β_f in $S[\rho(\beta)]$. For the bath this difference should not be neglected, since this small difference is multiplied by the large number of the bath degrees of freedom. We write

$$S[\rho_B(\beta_f)] = S_B(U_B + \delta U_B), \tag{C9}$$

where U_B is the initial bath energy, and where δU_B is the change of the bath energy due to its interaction with the small system. Using $\mathrm{d}S_B/\mathrm{d}U_B = \beta$ and expanding

$$S_B(U_B + \delta U_B) = S_B(U_B) + \beta \delta U_B, \tag{C10}$$

we get from (C8): $\delta U_B = T(\,S[\rho] - S[\rho(\beta)]\,)$. Putting this into (C5) we get

$$W_{\mathrm{th}} = \mathrm{tr}\,[H\rho] - \mathrm{tr}\,[H\rho(\beta)] - T(\,S[\rho] - S[\rho(\beta)]\,) \tag{C11}$$
$$= F_i - F_f. \tag{C12}$$

This is the difference $F_i - F_f$ between the free energies, provided the latter is defined as

$$F = \mathrm{tr}\,[H\rho] - TS[\rho]. \tag{C13}$$

Note that W_{th} in (C12) can be written as $W_{\mathrm{th}} = T\,S[\,\rho\,||\,\rho(\beta)\,]$, where $S[\,\rho\,||\,\rho(\beta)\,]$ defined in (21), is the relative entropy between the initial state ρ and the final equilibrium state $\rho(\beta)$ of R.

2. Reachability of the thermodynamical upper bound via a slow process.

One may hope that the thermodynamical bound (C12) could indeed be reached by some realistic work extraction dynamics [1, 2, 21, 25]. This is because for a macroscopic system (for the present case R+B) the conservation of entropy alone is expected to characterize a thermally isolated process. More specifically, we should demand from a physical realization of the bound (C12) that the work-sources act on the system and, at worst, at the system bath-coupling, but not on the bath itself, since the latter is supposed to be out of a direct control.

A realization of the bound (C12) for restricted initial states ρ was outlined in [21]. This realization takes a long time and thus amounts to zero power of work-extraction. A finite power protocol of extracting (C12) from an arbitrary initial state ρ is given in [38]. However, this protocol has another serious drawback, namely one has to allow direct interactions between the source of work and the thermal bath (uncontrollable degrees of freedom).

The result presented in [21] focussed on the zero Hamiltonian case $H_{\rm R} = 0$. It is, however, possible to generalize this result assuming that the initial state ρ is diagonal in the representation of $H_{\rm R}$. This restriction is essential, as we discuss below. Thus we write $H_{\rm R}$ and ρ as

$$H = \mathrm{diag}[\varepsilon_n, ..., \varepsilon_1], \quad \varepsilon_n \geq ... \geq \varepsilon_1 = 0, \quad (\text{C14})$$
$$\rho = \mathrm{diag}[r_n, ..., r_1], \quad (\text{C15})$$

where $\mathrm{diag}[a, .., b]$ is a diagonal matrix with entries $(a, ..., b)$. For simplicity we assume that the eigenvalues r_k are ordered as

$$r_n \leq ... \leq r_1. \quad (\text{C16})$$

This assumption is not essential. If it does not hold, one should supplements the step **1** below by the unitary transformation that orders properly the spectrum of ρ.

The work extraction consists of the following two steps:

1. One changes with time the level spacings from their initial values $\{\varepsilon_n, ..., \varepsilon_2\}$ to final values $\{\varepsilon'_n, ..., \varepsilon'_2\}$:

$$\{\varepsilon_n, ..., \varepsilon_2\} \rightarrow \{\varepsilon'_n, ..., \varepsilon'_2\}. \quad (\text{C17})$$

The change occurs much *faster* than the relaxation time induced by the bath, so that the interaction with the bath can be neglected during the change. Since the corresponding time-dependent Hamiltonian commutes with the initial density matrix for all times, the populations $\{r_n, ..., r_1\}$ remain constant during this process. The purpose of (C17) is to achieve the *local equilibrium* of the system at the bath temperature β:

$$r_n = \frac{e^{-\beta \varepsilon'_n}}{Z}, \quad ..., \quad r_2 = \frac{e^{-\beta \varepsilon'_2}}{Z}, \quad r_1 = \frac{1}{Z}, \quad (\text{C18})$$

where $Z \equiv \sum_{k=1}^{n} e^{-\beta \varepsilon'_k}$. Together with (C14, C16), Eqs. (C18) define the new spacings $\{\varepsilon'_n, ..., \varepsilon'_2\}$:

$$\varepsilon'_n = T_{\rm i} \ln \frac{r_1}{r_n}, ..., \quad \varepsilon'_n = T_{\rm i} \ln \frac{r_1}{r_2}. \quad (\text{C19})$$

The full work W_1 done during (C17) is [initial average energy minus final average energy]

$$W_1 = \sum_{k=2}^{n} r_k [\varepsilon_k - \varepsilon'_k]. \quad (\text{C20})$$

2. The spacings $\varepsilon'_n, ..., \varepsilon'_2$ are now slowly moved back to their original values $\varepsilon_n, ..., \varepsilon_2$. Here slow means that the characteristic time of the variation is much larger than the relaxation time of the system (determined by the coupling to the bath). It is at this point that the work-extraction process is going to take a long time.

During this process the density matrix is

$$\rho_2(t) = \frac{1}{Z(t)} \mathrm{diag} \left[e^{-\beta \varepsilon_n(t)}, \quad ... \quad , e^{-\beta \varepsilon_2(t)}, 1 \right] (\text{C21})$$

$$Z(t) \equiv \sum_{k=1}^{n} e^{-\beta \varepsilon_k(t)}, \quad (\text{C22})$$

where at the initial stage of this second step $\varepsilon_k(t_{\rm i}) = \varepsilon'_k$, while at the final stage $\varepsilon_k(t_{\rm f}) = \varepsilon_k$.

The work done during this process reads

$$W_2 = -\int_{t_{\rm i}}^{t_{\rm f}} \mathrm{d}s \, \rho_2(s) \partial_s H(s) = \int_{t_{\rm i}}^{t_{\rm f}} \mathrm{d}s \, \frac{\mathrm{d}\left[T \ln Z(s) \right]}{\mathrm{d}s}, \quad (\text{C23})$$

where $Z(t)$ is defined in (C22). Working out (C23) one gets that W_2 is equal to the free energy difference:

$$W_2 = -\left[\mathrm{tr}(\rho(\beta)H) - T S[\rho(\beta)] \right] \quad (\text{C24})$$

$$+ \sum_{k=2}^{n} r_k \varepsilon'_k + T \sum_{k=1}^{n} r_k \ln r_k, \quad (\text{C25})$$

where the RHS of (C24) is the minus final free energy with $\rho(\beta) = \rho_2(t_{\rm f})$ being the final equilibrium state of R, and where (C25) is the free energy at the end of the sudden change defined in (C17). It is clear that the sum $W_1 + W_2$ defined via (C20, C24, C25) is equal to the thermodynamic bound (C12).

Since the local equilibrium is related to the simultaneously diagonal form for both H and ρ_0, would there be initial coherences [non-diagonal elements of ρ_0], there would be no unitary operation that could bring the system to the local equilibrium with the bath at the end of the first step. Thus the adopted restriction on the initial state is essential.